

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 10:19:12 ; Search time 178.268 Seconds
(without alignments)
846.121 Million cell updates/sec

Title: US-10-053-975A-1
Perfect score: 2047
Sequence: 1 MAVSESQLKKMVSKYKYRDL.....FQLRALMQKARKTAGLSDLY 390

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	2047	100.0	390	2	AAW93424	Aaw93424 Human NHT
2	2047	100.0	390	5	ABP52191	Abp52191 Human tum
3	2047	100.0	390	6	ABG73787	Abg73787 Human tum
4	2047	100.0	390	7	ADC35176	Adc35176 Novel hum
5	2047	100.0	390	8	ADF91425	Adf91425 TSG101 #S
6	2047	100.0	390	8	ADL82957	Adl82957 Human PRO
7	2047	100.0	390	8	ADN12256	Adn12256 Human TSG
8	2047	100.0	390	8	ADO43235	Ado43235 Human TSG
9	2047	100.0	390	8	ADO43233	Ado43233 Human TSG

10	2047	100.0	390	8	ADP82614	Adp82614	Human	tum
11	2047	100.0	390	8	ADP56114	Adp56114	Human	PRO
12	2042	99.8	390	6	AAE34885	Aae34885	Human	Tsg
13	2002	97.8	380	2	AAW19111	Aaw19111	Human	tum
14	1945.5	95.0	391	6	AAE34884	Aae34884	Mouse	Tsg
15	1945.5	95.0	391	7	ADB85228	Adb85228	Mouse	tum
16	1900.5	92.8	381	2	AAW19110	Aaw19110	Mouse	tum
17	1900.5	92.8	381	2	AAW93425	Aaw93425	Mouse	tsg
18	1900.5	92.8	381	6	ABG73788	Abg73788	Murine	ts
19	1900.5	92.8	381	7	ADC35178	Adc35178	Human	tsg
20	1404	68.6	307	5	ABP41729	Abp41729	Human	ova
21	1162	56.8	237	4	AAE09328	Aae09328	Human	int
22	809.5	39.5	331	4	ABB64607	Abb64607	Drosophil	
23	776	37.9	145	8	ADF91424	Adf91424	TSG101UEV	
24	479	23.4	90	3	AAG01689	Aag01689	Human	sec
25	455	22.2	398	3	AAG06370	Aag06370	Arabidops	
26	455	22.2	398	8	ADN73283	Adn73283	Thale cre	
27	455	22.2	412	3	AAG06369	Aag06369	Arabidops	
28	435	21.3	379	5	ABP43551	Abp43551	L-lactate	
29	435	21.3	379	8	ADH13697	Adh13697	Human	ENZ
30	424	20.7	340	3	AAG06371	Aag06371	Arabidops	
31	390	19.1	87	8	ABO55273	Abo55273	Human	gen
32	380	18.6	322	3	AAG13898	Aag13898	Arabidops	
33	371	18.1	341	4	AAB93473	Aab93473	Human	pro
34	371	18.1	433	7	ADM29340	Adm29340	Human	nov
35	346	16.9	146	3	AAB53717	Aab53717	Human	col
36	339	16.6	452	4	AAU23116	Aau23116	Novel	hum
37	328	16.0	73	8	ABO55495	Abo55495	Human	gen
38	267.5	13.1	288	3	AAG13899	Aag13899	Arabidops	
39	161.5	7.9	168	3	AAG13900	Aag13900	Arabidops	
40	161	7.9	485	3	AAB57086	Aab57086	Human	pro
41	160	7.8	466	6	ABO07215	Abo07215	Human	p53
42	160	7.8	466	7	ADD47589	Add47589	Human	Pro
43	160	7.8	466	7	ADI15879	Adi15879	Human	PP
44	160	7.8	466	7	ADJ68571	Adj68571	Human	hea
45	153	7.5	218	4	ABG18833	Abg18833	Novel	hum
46	151	7.4	466	2	AAR07084	Aar07084	Recombina	
47	148.5	7.3	148	3	AAy86515	Aay86515	Human	gen
48	148.5	7.3	148	5	ABB97431	Abb97431	Novel	hum
49	148.5	7.3	148	6	ABO53679	Abo53679	Novel	hum
50	148.5	7.3	148	8	ABM81395	Abm81395	Tumour-as	
51	148.5	7.3	176	4	AAG75132	Aag75132	Human	col
52	148.5	7.3	397	4	AAM39852	Aam39852	Human	pol
53	148.5	7.3	397	4	AAG81381	Aag81381	Human	AFP
54	148.5	7.3	397	6	ADA55415	Ada55415	Human	pro
55	148.5	7.3	442	4	AAM41638	Aam41638	Human	pol
56	145.5	7.1	368	8	ADM97711	Adm97711	Human	MNA
57	145.5	7.1	536	8	ADM97710	Adm97710	Human	MNA
58	145.5	7.1	1130	8	ADM97695	Adm97695	Human	MNA
59	144.5	7.1	777	5	ABP41839	Abp41839	Human	ova
60	144.5	7.1	1130	7	ADC99065	Adc99065	Human	KPP
61	144.5	7.1	1130	8	ABM81812	Abm81812	Tumour-as	
62	142	6.9	505	2	AAy07117	Aay07117	Lung	canc
63	142	6.9	505	6	ABO07216	Abo07216	Human	p53
64	142	6.9	505	7	ADJ68733	Adj68733	Human	hea
65	142	6.9	505	8	ABM81156	Abm81156	Tumour-as	
66	142	6.9	505	8	ADP23965	Adp23965	PRO	polyp

67	142	6.9	905	2	AAW31186	Aaw31186	Human	p16
68	142	6.9	983	8	ADM87342	Adm87342	Human	pro
69	142	6.9	1021	7	ADJ68507	Adj68507	Human	hea
70	142	6.9	1135	2	AAW31185	Aaw31185	Human	p16
71	142	6.9	1251	7	ADE09332	Ade09332	Novel	pro
72	142	6.9	1251	8	ADM87788	Adm87788	Human	EST
73	141	6.9	244	3	AAG40015	Aag40015	Arabidops	
74	141	6.9	245	3	AAG38059	Aag38059	Arabidops	
75	141	6.9	245	3	AAG05889	Aag05889	Arabidops	
76	141	6.9	247	3	AAG38058	Aag38058	Arabidops	
77	141	6.9	247	3	AAG05888	Aag05888	Arabidops	
78	141	6.9	263	3	AAG38057	Aag38057	Arabidops	
79	141	6.9	274	3	AAG05887	Aag05887	Arabidops	
80	141	6.9	463	7	ADD47587	Add47587	Rat	Prote
81	140.5	6.9	1277	4	ABB66688	Abb66688	Drosophil	
82	139.5	6.8	258	3	AAB44500	Aab44500	Plant	vir
83	138.5	6.8	827	5	ABP35602	Abp35602	Fungal	ZB
84	135.5	6.6	177	4	ABG09342	Abg09342	Novel	hum
85	135	6.6	121	3	AAG21511	Aag21511	Arabidops	
86	134.5	6.6	172	4	ABG18832	Abg18832	Novel	hum
87	133	6.5	144	4	AAM42061	Aam42061	Human	pol
88	132	6.4	369	7	ADD14178	Add14178	Human	src
89	132	6.4	371	2	AAW64535	Aaw64535	Human	leu
90	132	6.4	371	5	ABP61444	Abp61444	Human	NF-
91	132	6.4	371	6	ABR41017	Abr41017	Human	MAP
92	132	6.4	371	8	ADR09413	Adr09413	Human	pro
93	131	6.4	101	3	AAG21512	Aag21512	Arabidops	
94	131	6.4	518	8	ADN20587	Adn20587	Bacterial	
95	131	6.4	539	3	AAB10224	Aab10224	Chicken	n
96	130.5	6.4	237	6	ABR58292	Abr58292	BCU0371B	
97	130.5	6.4	338	7	ADB75310	Adb75310	Prostate	
98	130.5	6.4	338	8	ADK70560	Adk70560	Respirato	
99	130.5	6.4	338	8	ADQ17947	Adq17947	Human	sof
100	130.5	6.4	338	8	ADP23414	Adp23414	PRO	polyp

ALIGNMENTS

RESULT 1

AAW93424

ID AAW93424 standard; protein; 390 AA.

XX

AC AAW93424;

XX

DT 11-JUN-1999 (first entry)

XX

DE Human NHTS protein.

XX

KW Human tumour suppressor; NHTS; treatment; diagnosis; lymphoma; cancer;
KW brain; breast; colon; heart; kidney; ovary; paraganglia; pancreas;
KW prostate; skin; stomach; thyroid; autoimmune disease; asthma; diabetes;
KW biliary cirrhosis; Crohn's disease; rheumatoid arthritis.

XX

OS Homo sapiens.

XX

PN US5892016-A.

XX PD 06-APR-1999.
XX
XX PF 23-JAN-1997; 97US-00786999.
XX
XX PR 23-JAN-1997; 97US-00786999.
XX
XX PA (INCY-) INCYTE PHARM.
XX
XX PI Brie SL, Goli SK;
XX
XX DR WPI; 1999-253932/21.
XX DR N-PSDB; AAX23168.
XX
XX PT Novel human tumor suppressor - useful for the diagnosis or treatment of
XX PT lymphoma, cancer, and autoimmune disease.
XX
XX PS Claim 1; Fig 1A-B; 25pp; English.
XX
XX CC This invention describes the isolation of a novel human tumour suppressor
XX CC (NHTS). The products of the invention may be used for the diagnosis or
XX CC treatment of conditions and diseases which are associated with expression
XX CC of NHTS e.g. lymphoma and cancers of the brain, breast, colon, heart,
XX CC kidney, ovary, paraganglia, pancreas, prostate, skin, stomach and thyroid
XX CC and autoimmune disease e.g. asthma, biliary cirrhosis, Crohn's disease,
XX CC diabetes, and rheumatoid arthritis
XX
XX SQ Sequence 390 AA;

Qy	1	MAVSESQLKKMVS	SKYKYRDLTVRET	VNVITLYKDLKPV	LD	SYVFNDGSSREL	MNLTGTIP	60
Db	1	MAVSESQLKKMVS	SKYKYRDLTVRET	VNVITLYKDLKPV	LD	SYVFNDGSSREL	MNLTGTIP	60
Qy	61	VPYRGNTYNIPIC	LWLLDTPYNPPIC	CFVKPTSSMTIKT	GKHVDANGKIYLP	YLHEWKHP		120
Db	61	VPYRGNTYNIPIC	LWLLDTPYNPPIC	CFVKPTSSMTIKT	GKHVDANGKIYLP	YLHEWKHP		120
Qy	121	QSDLLGLIQVMIV	VFGDEPPVFSRP	ISASYP	PPYQATGPPNTS	SYMPGMPGGIS	PYPSGYPP	180
Db	121	QSDLLGLIQVMIV	VFGDEPPVFSRP	ISASYP	PPYQATGPPNTS	SYMPGMPGGIS	PYPSGYPP	180
Qy	181	NPSGYPGCPYPPG	GPYPATTSSQYPS	QPPVTTVGPSRD	GTISED	TIRASLISAVS	DKLRW	240
Db	181	NPSGYPGCPYPPG	GPYPATTSSQYPS	QPPVTTVGPSRD	GTISED	TIRASLISAVS	DKLRW	240
Qy	241	RMKEEMDRAQAE	LNALKRTEEDLKK	GHQKLEEMVTRL	DQEVAEVDKNI	ELLKKKDEELSS		300
Db	241	RMKEEMDRAQAE	LNALKRTEEDLKK	GHQKLEEMVTRL	DQEVAEVDKNI	ELLKKKDEELSS		300
Qy	301	ALEKMENQSENND	IDEVIIPTAPLYK	QIILNLYAEENA	IEDTIFYLGEAL	R	RGVIDLDVFL	360
Db	301	ALEKMENQSENND	IDEVIIPTAPLYK	QIILNLYAEENA	IEDTIFYLGEAL	R	RGVIDLDVFL	360

Qy 361 KHVRLLSRKQFQLRALMQKARKTAGLSLDLY 390
 |||
 Db 361 KHVRLLSRKQFQLRALMQKARKTAGLSLDLY 390

us-10-053-975a-1.ra1

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 10:33:17 ; Search time 46.8487 Seconds
(without alignments)
621.429 Million cell updates/sec

Title: US-10-053-975A-1
Perfect score: 2047
Sequence: 1 MAVSESQLKKMVSKYKYRDL.....FQLRALMQKARKTAGLSDLY 390

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2047	100.0	390	2	US-08-786-999-1	Sequence 1, Appli
2	2047	100.0	390	4	US-09-216-387-1	Sequence 1, Appli
3	2047	100.0	390	4	US-09-886-319A-2	Sequence 2, Appli
4	2047	100.0	403	4	US-09-949-016-11251	Sequence 11251, A
5	2002	97.8	380	1	US-08-585-758A-4	Sequence 4, Appli
6	2002	97.8	380	1	US-08-977-818-4	Sequence 4, Appli
7	2002	97.8	380	2	US-08-670-274B-4	Sequence 4, Appli
8	2002	97.8	380	3	US-09-146-187-4	Sequence 4, Appli
9	2002	97.8	380	4	US-09-804-690-4	Sequence 4, Appli
10	1945.5	95.0	391	4	US-09-886-319A-1	Sequence 1, Appli
11	1900.5	92.8	381	1	US-08-585-758A-2	Sequence 2, Appli
12	1900.5	92.8	381	1	US-08-977-818-2	Sequence 2, Appli
13	1900.5	92.8	381	2	US-08-670-274B-2	Sequence 2, Appli
14	1900.5	92.8	381	2	US-08-786-999-3	Sequence 3, Appli
15	1900.5	92.8	381	3	US-09-146-187-2	Sequence 2, Appli
16	1900.5	92.8	381	4	US-09-216-387-3	Sequence 3, Appli
17	1900.5	92.8	381	4	US-09-804-690-2	Sequence 2, Appli

				us-10-053-975a-1.ra		
18	1404.5	68.6	285	4	US-09-886-319A-82	Sequence 82, Appl
19	1162	56.8	237	3	US-08-999-774A-8	Sequence 8, Appli
20	852.5	41.6	369	4	US-09-270-767-43042	Sequence 43042, A
21	479	23.4	90	4	US-09-513-999C-5770	Sequence 5770, Ap
22	342.5	16.7	173	4	US-09-270-767-58371	Sequence 58371, A
23	173	8.5	361	4	US-09-248-796A-16547	Sequence 16547, A
24	160	7.8	466	3	US-08-526-136-13	Sequence 13, Appl
25	160	7.8	467	4	US-09-949-016-7070	Sequence 7070, Ap
26	160	7.8	467	4	US-09-949-016-7071	Sequence 7071, Ap
27	148.5	7.3	148	4	US-09-461-325-453	Sequence 453, App
28	148.5	7.3	148	4	US-10-012-542-453	Sequence 453, App
29	148.5	7.3	148	4	US-10-115-123-453	Sequence 453, App
30	147	7.2	489	4	US-09-949-016-7068	Sequence 7068, Ap
31	147	7.2	489	4	US-09-949-016-7069	Sequence 7069, Ap
32	142	6.9	557	4	US-09-949-016-7621	Sequence 7621, Ap
33	141	6.9	905	2	US-08-574-959A-9	Sequence 9, Appli
34	141	6.9	905	3	US-09-357-014-9	Sequence 9, Appli
35	141	6.9	1135	2	US-08-574-959A-7	Sequence 7, Appli
36	141	6.9	1135	3	US-09-357-014-7	Sequence 7, Appli
37	130.5	6.4	452	4	US-09-949-016-7289	Sequence 7289, Ap
38	130.5	6.4	475	4	US-09-248-796A-14568	Sequence 14568, A
39	130	6.4	543	4	US-09-535-008-63	Sequence 63, Appl
40	130	6.4	577	4	US-09-535-008-61	Sequence 61, Appl
41	130	6.4	1646	4	US-09-535-008-67	Sequence 67, Appl
42	130	6.4	1647	4	US-09-535-008-2	Sequence 2, Appli
43	130	6.4	1647	4	US-09-824-574-4	Sequence 4, Appli
44	130	6.4	1647	4	US-09-538-092-1172	Sequence 1172, Ap
45	130	6.4	1649	4	US-09-535-008-75	Sequence 75, Appl
46	130	6.4	1650	4	US-09-535-008-71	Sequence 71, Appl
47	130	6.4	1659	4	US-09-949-016-9752	Sequence 9752, Ap
48	130	6.4	1678	4	US-09-535-008-69	Sequence 69, Appl
49	130	6.4	1679	4	US-09-535-008-65	Sequence 65, Appl
50	130	6.4	1681	4	US-09-535-008-77	Sequence 77, Appl
51	130	6.4	1682	4	US-09-535-008-73	Sequence 73, Appl
52	127	6.2	1184	4	US-09-266-225D-18	Sequence 18, Appl
53	127	6.2	1185	3	US-09-041-886-23	Sequence 23, Appl
54	127	6.2	1185	4	US-09-538-092-1209	Sequence 1209, Ap
55	126	6.2	172	4	US-09-800-729-138	Sequence 138, App
56	126	6.2	868	4	US-09-800-729-106	Sequence 106, App
57	126	6.2	921	4	US-09-800-729-199	Sequence 199, App
58	126	6.2	8991	4	US-08-714-741-32	Sequence 32, Appl
59	125.5	6.1	330	1	US-08-642-255-32	Sequence 32, Appl
60	125.5	6.1	408	1	US-07-609-716-65	Sequence 65, Appl
61	125.5	6.1	408	3	US-08-475-411A-65	Sequence 65, Appl
62	125.5	6.1	408	3	US-08-478-029A-65	Sequence 65, Appl
63	124.5	6.1	1162	2	US-08-728-323A-2	Sequence 2, Appli
64	124.5	6.1	1162	3	US-09-298-568-2	Sequence 2, Appli
65	124.5	6.1	1162	4	US-09-410-399-2	Sequence 2, Appli
66	124.5	6.1	1162	4	US-09-894-273-2	Sequence 2, Appli
67	124.5	6.1	1274	3	US-09-095-443-2	Sequence 2, Appli
68	123.5	6.0	258	4	US-09-270-767-46446	Sequence 46446, A
69	123.5	6.0	380	2	US-09-026-587-4	Sequence 4, Appli
70	123.5	6.0	380	2	US-09-227-420-4	Sequence 4, Appli
71	123.5	6.0	380	4	US-09-387-811-4	Sequence 4, Appli
72	123.5	6.0	380	4	US-09-823-240A-10	Sequence 10, Appl
73	123.5	6.0	381	4	US-09-949-016-7291	Sequence 7291, Ap
74	123.5	6.0	635	4	US-09-417-197-125	Sequence 125, App
75	123	6.0	1248	2	US-09-080-897-2	Sequence 2, Appli
76	123	6.0	1248	3	US-09-323-735-2	Sequence 2, Appli
77	122	6.0	1315	3	US-08-899-595-3	Sequence 3, Appli
78	121.5	5.9	420	4	US-09-902-540-13993	Sequence 13993, A
79	121.5	5.9	553	4	US-09-949-016-7961	Sequence 7961, Ap
80	120	5.9	777	4	US-09-949-016-8706	Sequence 8706, Ap

					us-10-053-975a-1.ra1	
81	120	5.9	777	4	US-09-949-016-8707	Sequence 8707, Ap
82	119.5	5.8	503	3	US-08-526-136-2	Sequence 2, Appli
83	119.5	5.8	505	3	US-08-526-136-4	Sequence 4, Appli
84	118.5	5.8	397	4	US-09-248-796A-18882	Sequence 18882, A
85	118.5	5.8	1255	2	US-09-080-897-4	Sequence 4, Appli
86	118.5	5.8	1255	3	US-08-899-595-1	Sequence 1, Appli
87	118.5	5.8	1255	3	US-09-323-735-4	Sequence 4, Appli
88	118	5.8	214	1	US-08-217-327-4	Sequence 4, Appli
89	118	5.8	581	4	US-09-949-016-9978	Sequence 9978, Ap
90	117	5.7	416	4	US-09-690-454-136	Sequence 136, App
91	117	5.7	779	4	US-10-164-595-56	Sequence 56, Appl
92	117	5.7	843	4	US-10-164-595-54	Sequence 54, Appl
93	116.5	5.7	338	1	US-08-218-686-2	Sequence 2, Appli
94	116.5	5.7	338	3	US-08-460-242-2	Sequence 2, Appli
95	116	5.7	1461	4	US-09-585-887-9	Sequence 9, Appli
96	116	5.7	1461	4	US-09-289-578-9	Sequence 9, Appli
97	116	5.7	1464	4	US-09-331-347C-21	Sequence 21, Appl
98	115.5	5.6	357	1	US-07-609-716-66	Sequence 66, Appl
99	115.5	5.6	357	1	US-08-642-255-33	Sequence 33, Appl
100	115.5	5.6	357	3	US-08-475-411A-66	Sequence 66, Appl

ALIGNMENTS

RESULT 1

US-08-786-999-1

; Sequence 1, Application US/08786999

; Patent No. 5892016

GENERAL INFORMATION:

APPLICANT: La Brie, Sam

APPLICANT: Goli, Surya K.

TITLE OF INVENTION: NOVEL HUMAN TUMOR

TITLE OF INVENTION: SUPPRESSOR

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSEQ for windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/786,999

FILING DATE: Filed Herewith

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0199 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

us-10-053-975a-1.ra1

: LENGTH: 390 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: Colnnot01
: CLONE: 609476
US-08-786-999-1

Query Match 100.0%; Score 2047; DB 2; Length 390;
Best Local Similarity 100.0%; Pred. No. 7.3e-159;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAVSESQLKKMVS	KYK	YRDL	TVRET	VNVIT	LYKDL	KPVLD	SYVFN	DGSSRE	LMNLT	GTIP	60
Db	1	MAVSESQLKKMVS	KYK	YRDL	TVRET	VNVIT	LYKDL	KPVLD	SYVFN	DGSSRE	LMNLT	GTIP	60
Qy	61	VPYRGNTYNIPIC	LWLLD	TPYNPP	ICFVK	PTSSMT	IKTGKH	VDANG	KIYLP	PYLHEW	KHP		120
Db	61	VPYRGNTYNIPIC	LWLLD	TPYNPP	ICFVK	PTSSMT	IKTGKH	VDANG	KIYLP	PYLHEW	KHP		120
Qy	121	QSDLLGLIQVMIV	VFGDE	PPVFSR	PISAS	YPPYQ	ATGPP	NTSYMP	GMPPG	GISPYP	SGYPP		180
Db	121	QSDLLGLIQVMIV	VFGDE	PPVFSR	PISAS	YPPYQ	ATGPP	NTSYMP	GMPPG	GISPYP	SGYPP		180
Qy	181	NPSGYPGCPYPPG	GYPAT	TSSQY	PSQPP	VTTVG	PSRDG	TISED	TIRAS	LISAVS	DKLRW		240
Db	181	NPSGYPGCPYPPG	GYPAT	TSSQY	PSQPP	VTTVG	PSRDG	TISED	TIRAS	LISAVS	DKLRW		240
Qy	241	RMKEEMDRAQAEL	NALKR	TEEDL	KKGHQ	KLEEM	VTRLD	QEVAE	VDKNI	ELLKKK	DEELSS		300
Db	241	RMKEEMDRAQAEL	NALKR	TEEDL	KKGHQ	KLEEM	VTRLD	QEVAE	VDKNI	ELLKKK	DEELSS		300
Qy	301	ALEKMQSENNDI	DEVII	PTAPL	YKQIL	NLYAE	ENAIED	TIFYL	GEALR	RGVIDL	DVFL		360
Db	301	ALEKMQSENNDI	DEVII	PTAPL	YKQIL	NLYAE	ENAIED	TIFYL	GEALR	RGVIDL	DVFL		360
Qy	361	KHVRLLSRKQFQ	LRLMQ	KARKT	AGLS	DLY							390
Db	361	KHVRLLSRKQFQ	LRLMQ	KARKT	AGLS	DLY							390

us-10-053-975a-1.rapb

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 10:46:54 ; Search time 160.016 seconds
(without alignments)
934.287 Million cell updates/sec

Title: US-10-053-975A-1
Perfect score: 2047
Sequence: 1 MAVSESQLKKMVKYKYRDL.....FQLRALMQKARKTAGLSDLY 390

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2047	100.0	390	9	US-09-886-319A-2	Sequence 2, Appli

				us-10-053-975a-1.rapb		
2	2047	100.0	390	14	US-10-243-815A-1	Sequence 1, Appli
3	2047	100.0	390	14	US-10-053-975A-1	Sequence 1, Appli
4	2047	100.0	390	14	US-10-376-564-2	Sequence 2, Appli
5	2047	100.0	390	16	US-10-675-979-1	Sequence 1, Appli
6	2002	97.8	380	9	US-09-804-690-4	Sequence 4, Appli
7	1945.5	95.0	391	9	US-09-886-319A-1	Sequence 1, Appli
8	1945.5	95.0	391	14	US-10-205-194-109	Sequence 109, App
9	1945.5	95.0	391	14	US-10-376-564-1	Sequence 1, Appli
10	1900.5	92.8	381	9	US-09-804-690-2	Sequence 2, Appli
11	1900.5	92.8	381	14	US-10-243-815A-3	Sequence 3, Appli
12	1404.5	68.6	285	9	US-09-886-319A-82	Sequence 82, Appl
13	1404.5	68.6	285	14	US-10-376-564-82	Sequence 82, Appl
14	1404	68.6	307	15	US-10-264-049-2861	Sequence 2861, Ap
15	396.5	19.4	402	16	US-10-437-963-164064	Sequence 164064,
16	390	19.1	87	14	US-10-029-386-28907	Sequence 28907, A
17	346	16.9	146	9	US-09-925-299-1257	Sequence 1257, Ap
18	346	16.9	146	10	US-09-925-299-1257	Sequence 1257, Ap
19	328	16.0	73	14	US-10-029-386-29129	Sequence 29129, A
20	269	13.1	246	16	US-10-425-115-299224	Sequence 299224,
21	266	13.0	219	16	US-10-425-115-295478	Sequence 295478,
22	231	11.3	231	15	US-10-424-599-239306	Sequence 239306,
23	217.5	10.6	266	16	US-10-425-115-298252	Sequence 298252,
24	200	9.8	296	9	US-09-801-368-386	Sequence 386, App
25	169.5	8.3	191	15	US-10-424-599-283387	Sequence 283387,
26	162	7.9	466	17	US-10-734-049A-223	Sequence 223, App
27	161	7.9	485	9	US-09-925-300-1664	Sequence 1664, Ap
28	160	7.8	466	16	US-10-408-765A-377	Sequence 377, App
29	148.5	7.3	148	14	US-10-012-542-453	Sequence 453, App
30	148.5	7.3	148	14	US-10-115-123-453	Sequence 453, App
31	148.5	7.3	176	14	US-10-106-698-5906	Sequence 5906, Ap
32	148.5	7.3	397	15	US-10-094-749-2983	Sequence 2983, Ap
33	145.5	7.1	1130	17	US-10-927-644-1	Sequence 1, Appli
34	145	7.1	483	16	US-10-437-963-191498	Sequence 191498,
35	144.5	7.1	777	15	US-10-264-049-2971	Sequence 2971, Ap
36	142.5	7.0	587	16	US-10-437-963-106668	Sequence 106668,
37	142	6.9	263	16	US-10-739-930-7139	Sequence 7139, Ap
38	142	6.9	505	16	US-10-408-765A-539	Sequence 539, App
39	142	6.9	505	16	US-10-719-993-624	Sequence 624, App
40	142	6.9	505	16	US-10-719-993-626	Sequence 626, App
41	142	6.9	505	17	US-10-952-459-10	Sequence 10, Appl
42	142	6.9	983	15	US-10-112-944-435	Sequence 435, App
43	142	6.9	1021	16	US-10-408-765A-313	Sequence 313, App
44	142	6.9	1251	15	US-10-112-944-881	Sequence 881, App
45	139.5	6.8	258	15	US-10-424-599-176931	Sequence 176931,
46	139.5	6.8	278	15	US-10-425-114-68284	Sequence 68284, A
47	138.5	6.8	827	15	US-10-149-310-68	Sequence 68, Appl
48	136	6.6	1066	16	US-10-437-963-157156	Sequence 157156,
49	135.5	6.6	180	16	US-10-437-963-191496	Sequence 191496,
50	134	6.5	1613	17	US-10-732-923-8825	Sequence 8825, Ap
51	133.5	6.5	336	16	US-10-437-963-201912	Sequence 201912,
52	132	6.4	371	10	US-09-284-320-2	Sequence 2, Appli
53	132	6.4	371	14	US-10-197-666A-4	Sequence 4, Appli
54	132	6.4	371	14	US-10-024-298A-41	Sequence 41, Appl
55	132	6.4	371	14	US-10-042-211A-41	Sequence 41, Appl
56	132	6.4	371	15	US-10-617-217A-41	Sequence 41, Appl
57	132	6.4	371	16	US-10-024-298A-41	Sequence 41, Appl
58	131.5	6.4	605	16	US-10-437-963-102875	Sequence 102875,
59	131.5	6.4	889	16	US-10-437-963-197045	Sequence 197045,
60	131	6.4	101	16	US-10-739-930-6173	Sequence 6173, Ap
61	131	6.4	101	17	US-10-732-923-2393	Sequence 2393, Ap
62	131	6.4	518	15	US-10-369-493-3240	Sequence 3240, Ap
63	131	6.4	588	16	US-10-437-963-150465	Sequence 150465,
64	131	6.4	873	16	US-10-437-963-142576	Sequence 142576,

				us-10-053-975a-1.rapb		
65	131	6.4	1613	17	US-10-732-923-8824	Sequence 8824, Ap
66	130.5	6.4	237	15	US-10-144-194A-30	Sequence 30, Appl
67	130.5	6.4	237	16	US-10-491-566-30	Sequence 30, Appl
68	130.5	6.4	338	14	US-10-205-823-134	Sequence 134, App
69	130.5	6.4	338	16	US-10-723-860-764	Sequence 764, App
70	130.5	6.4	1322	15	US-10-369-493-13040	Sequence 13040, A
71	130	6.4	1647	10	US-09-824-574-4	Sequence 4, Appli
72	130	6.4	1647	16	US-10-370-715B-356	Sequence 356, App
73	130	6.4	1647	17	US-10-732-923-8821	Sequence 8821, Ap
74	130	6.4	1647	17	US-10-732-923-8822	Sequence 8822, Ap
75	130	6.4	1679	17	US-10-732-923-8823	Sequence 8823, Ap
76	129.5	6.3	358	16	US-10-437-963-153141	Sequence 153141,
77	129.5	6.3	753	15	US-10-369-493-2333	Sequence 2333, Ap
78	129.5	6.3	1419	16	US-10-473-115-2	Sequence 2, Appli
79	129	6.3	1449	16	US-10-408-765A-1810	Sequence 1810, Ap
80	128.5	6.3	262	15	US-10-133-234A-5	Sequence 5, Appli
81	128.5	6.3	754	16	US-10-437-963-112229	Sequence 112229,
82	128	6.3	1121	16	US-10-437-963-196036	Sequence 196036,
83	127.5	6.2	256	16	US-10-425-115-187693	Sequence 187693,
84	127.5	6.2	284	13	US-10-006-867-62	Sequence 62, Appl
85	127.5	6.2	284	13	US-10-052-586-236	Sequence 236, App
86	127.5	6.2	284	13	US-10-063-547-62	Sequence 62, Appl
87	127.5	6.2	284	13	US-10-063-551-62	Sequence 62, Appl
88	127.5	6.2	284	14	US-10-174-590-236	Sequence 236, App
89	127.5	6.2	284	14	US-10-176-758-236	Sequence 236, App
90	127.5	6.2	284	14	US-10-175-737-236	Sequence 236, App
91	127.5	6.2	284	14	US-10-063-616-62	Sequence 62, Appl
92	127.5	6.2	284	14	US-10-174-581-236	Sequence 236, App
93	127.5	6.2	284	14	US-10-176-483-236	Sequence 236, App
94	127.5	6.2	284	14	US-10-176-749-236	Sequence 236, App
95	127.5	6.2	284	14	US-10-176-914-236	Sequence 236, App
96	127.5	6.2	284	14	US-10-176-915-236	Sequence 236, App
97	127.5	6.2	284	14	US-10-063-569-62	Sequence 62, Appl
98	127.5	6.2	284	14	US-10-063-513-62	Sequence 62, Appl
99	127.5	6.2	284	14	US-10-063-515-62	Sequence 62, Appl
100	127.5	6.2	284	14	US-10-063-512-62	Sequence 62, Appl

ALIGNMENTS

RESULT 1

US-09-886-319A-2

```

; Sequence 2, Application US/09886319A
; Publication No. US20020086019A1
; GENERAL INFORMATION:
; APPLICANT: Wolf, Eckard
; APPLICANT: Werner, Sabine
; APPLICANT: Halle, Jorn-Peter
; APPLICANT: Regenbogen, Johannes
; APPLICANT: Goppelt, Andreas
; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
; TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and wound
; TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
; TITLE OF INVENTION: Active Substances
; FILE REFERENCE: 50125/014002
; CURRENT APPLICATION NUMBER: US/09/886,319A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/222,081
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: DE 10030149.5
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 84

```

us-10-053-975a-1.rapb

; SOFTWARE: FastSEQ for windows Version 4.0

; SEQ ID NO 2

; LENGTH: 390

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-886-319A-2

Query Match 100.0%; Score 2047; DB 9; Length 390;
Best Local Similarity 100.0%; Pred. No. 3.2e-133;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MAVSESQLKKMVSKYKYRDLTVRETVNVITLYKDLKPVLD SYVFNDGSSRELMNLTGTIP 60
          |||
Db      1 MAVSESQLKKMVSKYKYRDLTVRETVNVITLYKDLKPVLD SYVFNDGSSRELMNLTGTIP 60

Qy     61 VPYRGNTYNIPICLWLLDTYPYNPPICFVKPTSSMTIKTGKHVDANGKIYLPYLHEWKHP 120
          |||
Db     61 VPYRGNTYNIPICLWLLDTYPYNPPICFVKPTSSMTIKTGKHVDANGKIYLPYLHEWKHP 120

Qy    121 QSDLLGLIQVMIVVFGDEPPVFSRPISASYPPYQATGPPNTSYMPGMPGGISPYPSGYPP 180
          |||
Db    121 QSDLLGLIQVMIVVFGDEPPVFSRPISASYPPYQATGPPNTSYMPGMPGGISPYPSGYPP 180

Qy    181 NPSGYPGCPYPPGGPYPATTSSQYPSQPPVTTVGPSRDGTISED TIRASLISAVSDKLRW 240
          |||
Db    181 NPSGYPGCPYPPGGPYPATTSSQYPSQPPVTTVGPSRDGTISED TIRASLISAVSDKLRW 240

Qy    241 RMKEEMDRAQAELNALKRTEEDLKKGHQKLEEMVTRL DQEVAEVDKNIELLKKKDEELSS 300
          |||
Db    241 RMKEEMDRAQAELNALKRTEEDLKKGHQKLEEMVTRL DQEVAEVDKNIELLKKKDEELSS 300

Qy    301 ALEKMENQSENNDIDEVIIPTAPLYKQILNLYAEENAIEDTIFYLGEALRRGVIDLDVFL 360
          |||
Db    301 ALEKMENQSENNDIDEVIIPTAPLYKQILNLYAEENAIEDTIFYLGEALRRGVIDLDVFL 360

Qy    361 KHVRLLSRKQFQLRALMQKARKTAGLS DLY 390
          |||
Db    361 KHVRLLSRKQFQLRALMQKARKTAGLS DLY 390
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 10:32:03 ; Search time 36.5055 Seconds
(without alignments)
1027.916 Million cell updates/sec

Title: US-10-053-975A-1
Perfect score: 2047
Sequence: 1 MAVSESQLKKMVSKYKYRDL.....FQLRALMQKARKTAGLSDLY 390

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : PIR_79:*

- 1: pirl:*
- 2: pir2:*
- 3: pir3:*
- 4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query				Description
		Match	Length	DB	ID	
1	658	32.1	404	2	H88650	protein C09G12.9 [
2	388	19.0	83	2	I48283	gene CC2 protein -
3	153.5	7.5	169	2	T34520	hypothetical prote
4	148.5	7.3	397	2	JC8036	hepatocellular car
5	147	7.2	488	1	LUHU7	annexin VII, long
6	145.5	7.1	212	2	S74288	hypothetical prote
7	142	6.9	505	2	A53152	annexin XI - human
8	141	6.9	463	2	S29170	annexin VII - mous
9	138.5	6.8	827	2	T39608	zinc finger transc
10	136.5	6.7	503	1	LURB11	annexin XI - rabbi
11	135.5	6.6	437	2	T14192	extensin homolog T
12	134	6.5	198	2	D70509	hypothetical prote
13	134	6.5	485	2	T37550	hypothetical coile

14	132.5	6.5	671	2	T36037	probable export as
15	131.5	6.4	370	2	T42532	hypothetical prote
16	131	6.4	1613	2	S39059	protein BRG1 - hum
17	131	6.4	1880	2	T18531	tractin - medicina
18	130.5	6.4	338	2	I53043	transforming prote
19	130	6.4	1647	2	S45252	SNF2beta protein -
20	128.5	6.3	1006	2	T42731	atrophin-1 related
21	128	6.3	139	2	H84809	hypothetical prote
22	128	6.3	3942	2	T42730	Bassoon protein -
23	127	6.2	1184	2	G01763	atrophin-1 - human
24	126.5	6.2	678	2	H88187	protein C18H9.8 [i
25	126	6.2	1453	2	S21626	collagen alpha 1(I
26	125	6.1	2715	2	T13049	eyelid - fruit fly
27	124.5	6.1	669	2	T28754	hypothetical prote
28	124.5	6.1	990	2	T14756	hypothetical prote
29	124	6.1	964	2	T21865	hypothetical prote
30	123.5	6.0	279	2	T05421	hypothetical prote
31	123.5	6.0	380	2	S51797	vasodilator-stimul
32	123	6.0	564	2	H70804	hypothetical prote
33	122	6.0	887	1	S57219	1-phosphatidylinos
34	122	6.0	978	2	A70387	conserved hypothet
35	122	6.0	1181	2	C86349	F8K7.4 protein - A
36	121.5	5.9	551	2	S57447	HPBRII-7 protein -
37	121	5.9	792	2	T49989	hypothetical prote
38	120.5	5.9	1902	2	C97702	cell surface antig
39	119.5	5.8	503	1	LUBO11	annexin XI form A
40	119.5	5.8	505	1	S23447	annexin XI form B
41	119	5.8	577	2	T09024	proline-rich prote
42	119	5.8	859	2	H70327	DNA mismatch repai
43	119	5.8	1357	2	T29265	hypothetical prote
44	119	5.8	1422	2	T24212	hypothetical prote
45	118.5	5.8	384	2	S51796	vasodilator-stimul
46	118.5	5.8	400	2	E70318	hypothetical prote
47	118.5	5.8	915	2	T12526	hypothetical prote
48	118.5	5.8	1255	2	T31065	diaphanous protein
49	118.5	5.8	1281	2	JC5368	dynactin 1 - mouse
50	118	5.8	176	2	A86441	hypothetical prote
51	118	5.8	214	2	T10737	extensin-like cell
52	118	5.8	214	2	T09854	proline-rich cell
53	118	5.8	530	2	T48627	hypothetical prote
54	118	5.8	1018	2	T43168	hypothetical prote
55	118	5.8	1171	2	T17454	diaphanous-related
56	118	5.8	1259	2	T16038	hypothetical prote
57	118	5.8	1736	2	T00391	hypothetical prote
58	117.5	5.7	316	2	T20497	hypothetical prote
59	117.5	5.7	620	2	S06733	hydroxyproline-ric
60	117.5	5.7	853	2	T51505	hypothetical prote
61	117.5	5.7	1802	2	H88444	protein C26E6.12 [
62	117	5.7	177	2	S65780	glycine/proline-ri
63	117	5.7	491	2	S14182	DNA-directed RNA p
64	117	5.7	505	2	S72273	actin-depolymerizi
65	117	5.7	650	2	S14181	DNA-directed RNA p
66	117	5.7	785	2	T09491	hemagglutinin, pha
67	117	5.7	1100	2	JC8033	leukocyte formin p
68	116.5	5.7	338	1	TVMSFB	transforming prote
69	116.5	5.7	1790	2	S67593	transport protein
70	116	5.7	393	2	PQ0479	pistil extensin-li

71	116	5.7	451	2	B70792	hypothetical prote
72	116	5.7	462	1	LUDO7	annexin VII - slim
73	116	5.7	504	2	A49467	occludin - chicken
74	116	5.7	513	2	T03916	hypothetical prote
75	116	5.7	609	2	S46019	YSW1 protein - yea
76	116	5.7	821	2	S67087	hypothetical prote
77	115.5	5.6	177	2	T47549	hypothetical prote
78	115.5	5.6	428	2	T24769	hypothetical prote
79	115.5	5.6	1189	2	T42726	guanine nucleotide
80	115	5.6	240	2	D70894	probable pra prote
81	114.5	5.6	240	2	A24264	proline-rich prote
82	114.5	5.6	272	2	T25608	hypothetical prote
83	114.5	5.6	324	2	S13497	cAMP-binding prote
84	114.5	5.6	398	2	T34947	hypothetical prote
85	114.5	5.6	554	2	F86244	hypothetical prote
86	114.5	5.6	817	2	S51342	verprolin - yeast
87	114.5	5.6	1633	2	JC5056	polybromo 1 - chic
88	114.5	5.6	1940	1	S04090	myosin heavy chain
89	114	5.6	342	2	A24263	myosin heavy chain
90	114	5.6	1017	2	T31354	probable potassium
91	114	5.6	1179	2	F71190	probable chromosom
92	114	5.6	1464	1	CGHU1S	collagen alpha 1(I
93	113.5	5.5	347	1	TVHUJB	transforming prote
94	113.5	5.5	1046	2	T42734	cytoplasmic linker
95	113.5	5.5	1487	1	CGHU6C	collagen alpha 1(I
96	113	5.5	242	2	S35060	tropomyosin - hydr
97	113	5.5	481	2	F86208	protein F22G5.30 [
98	113	5.5	1638	2	A42091	transcription acti
99	112.5	5.5	199	2	S14981	extensin class I (
100	112.5	5.5	324	2	G86222	hypothetical prote

ALIGNMENTS

RESULT 1

H88650

protein C09G12.9 [imported] - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C;Accession: H88650

R;anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology.

A;Reference number: A75000; MUID:99069613; PMID:9851916

A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and

www.sanger.ac.uk/Projects/C_elegans/ for a list of authors

A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 285, 1493, 1999

A;Accession: H88650

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-404 <STO>

A;Cross-references: UNIPROT:O76258; GB:chr_IV; PIDN:AAC25822.1; PID:g3294495;

GSPDB:GN00022; CESP:C09G12.9

C;Genetics:

A;Gene: C09G12.9
A;Map position: 4

Query Match 32.1%; Score 658; DB 2; Length 404;
Best Local Similarity 36.7%; Pred. No. 1.1e-30;
Matches 151; Conservative 74; Mismatches 121; Indels 66; Gaps 11;

```
Qy      16 KYRDLTVRETNNVITLYKDLKPVLD SYVFNDGSSREL MNLTGTIPV PYRGNTYNIPICLW 75
      || |  :: :  :: :||| | |:::| || |  | |||| |::| |||: ::
Db      16 KYADSAKKDIIGALSQFKDLSPGTDTFMFPDGKRRTAFLKGTIPVYYKGACYNIPVTVY 75

Qy      76 LLDTYPYNPPICFVKPTSSMTIKTGKHVDANGKIYLPYLHEWKHPQSDLLGLIQVMIVVF 135
      | ||:| | ||:| ||:|  :||: ||:||||:|:| | || |:|:
Db      76 LWDTHPYYPICYVNPSTSTM---ESEHVNKEGKVFLPYLNEWRFPGYDLSGLLQM----- 127

Qy     136 GDEPPVFSRPISASYPPYQATGP-----PNTSYMPGMPGGISPYPSG-----YP 179
      :|:| : |  || |  | | | |  :||:|  ||
Db     128 -----IFARSAANSATNASATNPSAGSSASSTPTPYPSQPTMP---TPYPTGSGAAPYP 179

Qy     180 PNPSGYPGC-----PYPPG----GPYPATTSSQY---PSQPPV---TTVG 214
      |: : ||  ||| |  |||: |:  | ||| |:|
Db     180 PSSTPYPSAGAMGYNPYMNVPQSTPYPMGASGASPYPSASSNPAPPPPRPPPVTAQTSVS 239

Qy     215 PSRDGTISED TIRASLISAVSDKLRWRMKEEMDRAQ AELNALKRTEEDLKKGHQKLEEMV 274
      | ||| |||||:|:| |:| :::| |  ||: :: | :|:|:| |||: |:
Db     240 SSSGGTIQADTIRASVMSAVEEKIRAKLRERMGTNSAEMASIRTTSD ELREGQQKLRML 299

Qy     275 TRLDQEVAEVDKNIELLKKKDEELSSALEKMENQSENNDIDEV IIPTAPLYKQILNLYAE 334
      |: : : :  |: | ||: ||  :: ||| |  ||:|:| |:|
Db     300 EELETQRSSLQTACEIYTAKKAELAKALSD-AGGTDAPPIDEAIDA AFPLHRQIVLNYAK 358

Qy     335 ENAIEDTIFYLGEALRRGVIDLDVFLKHVRLLSRKQFQLRALMQKARKTAGL 386
      :  :::: | | :|:| | :|:| | || ||| |:| |
Db     359 DLTC-----QSLKKRQITLAEYLRHVRDVSREQFIYRATMQKCRRTAGL 402
```

us-10-053-975a-1.rup

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 10:20:07 ; Search time 168.534 seconds
(without alignments)
1184.992 Million cell updates/sec

Title: US-10-053-975A-1
Perfect score: 2047
Sequence: 1 MAVSESQLKKMVSKEYKYRDL.....FQLRALMQKARKTAGLSLDLY 390

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2047	100.0	390	1	T101_HUMAN	Q99816 homo sapien
2	2041	99.7	390	2	Q9BUM5	Q9bum5 homo sapien
3	1945.5	95.0	391	1	T101_MOUSE	Q61187 mus musculu
4	1927.5	94.2	391	2	Q6IRE4	Q6ire4 rattus norv
5	1908.5	93.2	391	2	Q7TSE5	Q7tse5 rattus norv
6	1899	92.8	392	2	Q9I8G8	Q9i8g8 chelonia my
7	1705.5	83.3	390	2	Q6IQ70	Q6iq70 brachydanio
8	1685	82.3	394	2	Q6NUD5	Q6nud5 xenopus lae
9	1682.5	82.2	395	2	Q6P2Z8	Q6p2z8 xenopus tro
10	1024.5	50.0	408	2	Q9VVA7	Q9vva7 drosophila
11	976	47.7	402	2	Q7Q6B6	Q7q6b6 anopheles g
12	754.5	36.9	425	2	O76258	O76258 caenorhabdi
13	749.5	36.6	326	2	Q6DDX9	Q6ddx9 xenopus lae
14	599	29.3	249	2	Q8MQZ0	Q8mqz0 drosophila
15	522	25.5	114	2	Q7T2M2	Q7t2m2 ameiurus ne
16	455	22.2	398	2	Q9LHG8	Q9lhg8 arabidopsis
17	435	21.3	379	2	Q8IX04	Q8ix04 homo sapien
18	432	21.1	580	2	Q7S4R9	Q7s4r9 neurospora
19	429.5	21.0	249	2	Q7TQD3	Q7tqd3 mus musculu
20	427	20.9	174	2	O08761	O08761 mus musculu
21	415.5	20.3	476	2	Q66KB7	Q66kb7 xenopus tro

us-10-053-975a-1.rup

22	398	19.4	471	2	Q6DBY5	Q6dby5	brachydanio
23	396.5	19.4	402	2	Q6ESB7	Q6esb7	oryza sativ
24	391	19.1	368	2	Q9FFY6	Q9fffy6	arabidopsis
25	388	19.0	83	2	Q6LBE4	Q6lbe4	mus musculu
26	371	18.1	177	2	Q96FF5	Q96ff5	homo sapien
27	371	18.1	341	2	Q9NUX7	Q9nux7	homo sapien
28	341	16.7	357	2	Q6P2F0	Q6p2f0	homo sapien
29	303	14.8	385	1	ST22_YEAST	P25604	saccharomyc
30	301.5	14.7	378	2	Q6C148	Q6c148	yarrowia li
31	301.5	14.7	378	2	Q873M7	Q873m7	yarrowia li
32	294	14.4	111	2	Q8BU96	Q8bu96	mus musculu
33	294	14.4	468	2	Q6FS29	Q6fs29	candida gla
34	239	11.7	496	2	Q6BID5	Q6bid5	debaryomyce
35	215	10.5	445	2	Q75EU1	Q75eu1	ashbya goss
36	169	8.3	376	2	P78998	P78998	saccharomyc
37	155.5	7.6	299	2	Q6CS18	Q6cs18	kluyveromyc
38	154.5	7.5	397	2	Q6NW27	Q6nw27	homo sapien
39	153.5	7.5	148	2	Q6FIG7	Q6fig7	homo sapien
40	153	7.5	283	2	Q641Z8	Q641z8	rattus norv
41	152	7.4	1158	2	Q6GM32	Q6gm32	xenopus lae
42	151.5	7.4	485	2	Q804G3	Q804g3	brachydanio
43	148.5	7.3	148	2	Q6NXQ6	Q6nxq6	homo sapien
44	148.5	7.3	148	2	Q9NZ81	Q9nz81	homo sapien
45	148.5	7.3	397	2	Q8NEZ2	Q8nez2	homo sapien
46	148.5	7.3	397	2	Q96DL9	Q96dl9	homo sapien
47	147.5	7.2	263	1	Y737_DROME	P83475	drosophila
48	147	7.2	397	2	Q8CHS8	Q8chs8	mus musculu
49	145.5	7.1	1130	2	Q8IZL8	Q8izl8	homo sapien
50	144.5	7.1	743	2	Q9BU60	Q9bu60	homo sapien
51	144.5	7.1	1048	2	Q96FT1	Q96ft1	homo sapien
52	144.5	7.1	1130	2	Q6NTE6	Q6nte6	homo sapien
53	144	7.0	483	2	Q7T391	Q7t391	brachydanio
54	144	7.0	488	1	ANX7_HUMAN	P20073	homo sapien
55	143	7.0	664	2	Q6CDQ5	Q6cdq5	yarrowia li
56	142	6.9	505	1	ANXB_HUMAN	P50995	homo sapien
57	142	6.9	1021	2	O15451	O15451	homo sapien
58	141.5	6.9	284	2	Q6BGN1	Q6bgn1	debaryomyce
59	141.5	6.9	503	1	ANXB_MOUSE	P97384	mus musculu
60	141.5	6.9	503	2	Q921F1	Q921f1	mus musculu
61	141	6.9	247	2	Q945K9	Q945k9	arabidopsis
62	141	6.9	463	1	ANX7_MOUSE	Q07076	mus musculu
63	141	6.9	463	2	Q922A2	Q922a2	mus musculu
64	140.5	6.9	1167	2	Q640S7	Q640s7	xenopus tro
65	139.5	6.8	900	2	Q6PSU8	Q6psu8	arabidopsis
66	139	6.8	415	2	Q8CCV9	Q8ccv9	mus musculu
67	139	6.8	1284	2	O15450	O15450	homo sapien
68	138.5	6.8	827	2	O60130	O60130	schizosacch
69	138.5	6.8	838	2	Q871B7	Q871b7	neurospora
70	138.5	6.8	4903	1	MLL3_MOUSE	Q8brh4	mus musculu
71	138	6.7	526	2	Q804G4	Q804g4	brachydanio
72	137	6.7	463	2	Q8BP75	Q8bp75	mus musculu
73	136.5	6.7	503	1	ANXB_RABIT	P33477	oryctolagus
74	136	6.6	137	2	Q9CQJ5	Q9cqj5	m mus muscu
75	136	6.6	1066	2	Q7XE90	Q7xe90	oryza sativ
76	135.5	6.6	437	2	Q9STN0	Q9stn0	arabidopsis
77	135.5	6.6	895	2	Q9LIW7	Q9liw7	oryza sativ
78	135.5	6.6	998	2	Q6GNM2	Q6gnm2	xenopus lae
79	135.5	6.6	1105	2	Q8BMJ4	Q8bmj4	mus musculu
80	135	6.6	463	2	Q8VIN2	Q8vin2	rattus norv
81	135	6.6	463	2	Q6IRJ7	Q6irj7	rattus norv
82	135	6.6	674	2	Q18106	Q18106	caenorhabdi
83	135	6.6	703	2	Q7SCK8	Q7sck8	neurospora
84	134.5	6.6	809	2	Q7ZVN7	Q7zvn7	brachydanio

					us-10-053-975a-1.rup		
85	134	6.5	198	2	O86316	O86316	mycobacteri
86	134	6.5	201	2	Q7D8K1	Q7d8k1	mycobacteri
87	134	6.5	485	2	Q9UTP6	Q9utp6	schizosacch
88	134	6.5	502	2	Q641H0	Q641h0	xenopus lae
89	133	6.5	198	2	Q7U0B9	Q7u0b9	mycobacteri
90	133	6.5	548	2	Q6DDW4	Q6ddw4	xenopus lae
91	132.5	6.5	471	2	Q9LD31	Q9ld31	crypthecodi
92	132.5	6.5	512	2	Q6NRE3	Q6nre3	xenopus lae
93	132.5	6.5	671	2	Q9Z502	Q9z502	streptomyce
94	132	6.4	371	2	Q7Z429	Q7z429	homo sapien
95	131.5	6.4	900	2	Q94B77	Q94b77	arabidopsis
96	131.5	6.4	900	2	Q9FIU0	Q9fiu0	arabidopsis
97	131	6.4	101	2	Q8LCL8	Q8lcl8	arabidopsis
98	131	6.4	472	2	O59907	O59907	neurospora
99	131	6.4	472	2	Q6M9D4	Q6m9d4	neurospora
100	131	6.4	475	2	Q7RZ53	Q7rz53	neurospora

ALIGNMENTS

RESULT 1

T101_HUMAN

ID T101_HUMAN STANDARD; PRT; 390 AA.
AC Q99816;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Tumor susceptibility gene 101 protein.
GN Name=TSG101;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=97148696; PubMed=9019400; DOI=10.1016/S0092-8674(00)81866-8;
RA Li L., Li X., Francke U., Cohen S.N.;
RT "The TSG101 tumor susceptibility gene is located in chromosome 11 band
RT p15 and is mutated in human breast cancer.";
RL Cell 88:143-154(1997).
RN [2]
RP ERRATUM.
RX MEDLINE=99080983; PubMed=9867424; DOI=10.1016/S0092-8674(00)89342-3;
RA Li L., Francke U., Cohen S.N.;
RL Cell 93:660-660(1998).
RN [3]
RP ALTERNATIVE SPLICING (ISOFORM 2).
RX MEDLINE=98031856; PubMed=9366528; DOI=10.1038/sj.onc.1201591;
RA Gayther S.A., Barski P., Batley S.J., Li L., de Foy K.A., Cohen S.N.,
RA Ponder B.A., Caldas C.;
RT "Aberrant splicing of the TSG101 and FHIT genes occurs frequently in
RT multiple malignancies and in normal tissues and mimics alterations
RT previously described in tumours.";
RL Oncogene 15:2119-2126(1997).
RN [4]
RP ALTERNATIVE SPLICING.
RX MEDLINE=97384943; PubMed=9242438;
RA Lee M.P., Feinberg A.P.;
RT "Aberrant splicing but not mutations of TSG101 in human breast
RT cancer.";
RL Cancer Res. 57:3131-3134(1997).

RN [5]
 RP ALTERNATIVE SPLICING.
 RX MEDLINE=99054675; PubMed=9840940; DOI=10.1038/sj.onc.1202529;
 RA Wagner K.-U., Dierisseau P., Rucker E.B. III, Robinson G.W.,
 RA Hennighausen L.;
 RT "Genomic architecture and transcriptional activation of the mouse and
 RT human tumor susceptibility gene TSG101: common types of shorter
 RT transcripts are true alternative splice variants.";
 RL Oncogene 17:2761-2770(1998).
 RN [6]
 RP INTERACTION WITH DMAP1.
 RX MEDLINE=20347709; PubMed=10888872; DOI=10.1038/77023;
 RA Rountree M.R., Bachman K.E., Baylin S.B.;
 RT "DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at
 RT replication foci.";
 RL Nat. Genet. 25:269-277(2000).
 RN [7]
 RP STRUCTURE BY NMR OF 1-145.
 RX MEDLINE=22001147; PubMed=12006492; DOI=10.1093/emboj/21.10.2397;
 RA Pornillos O., Alam S.L., Rich R.L., Myszkka D.G., Davis D.R.,
 RA Sundquist W.I.;
 RT "Structure and functional interactions of the Tsg101 UEV domain.";
 RL EMBO J. 21:2397-2406(2002).
 RN [8]
 RP STRUCTURE BY NMR OF 1-145.
 RX MEDLINE=22289677; PubMed=12379843; DOI=10.1038/nsb856;
 RA Pornillos O., Alam S.L., Davis D.R., Sundquist W.I.;
 RT "Structure of the Tsg101 UEV domain in complex with the PTAP motif of
 RT the HIV-1 p6 protein.";
 RL Nat. Struct. Biol. 9:812-817(2002).
 CC -!- FUNCTION: May be involved in cell growth and differentiation. Acts
 CC as a negative growth regulator.
 CC -!- SUBUNIT: Interacts with ubiquitin, stathmin, GMCL and DMAP1 (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Mainly cytoplasmic. Depending on the stage
 CC of the cell cycle, detected in the nucleus.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Comment=Additional isoforms seem to exist. Several shorter
 CC isoforms are detected in primary breast cancers and other
 CC tumors;
 CC Name=1;
 CC IsoId=Q99816-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q99816-2; Sequence=VSP_004440;
 CC Note=Detected in normal as well as cancer tissues;
 CC -!- TISSUE SPECIFICITY: Heart, brain, placenta, lung, liver, skeletal,
 CC kidney and pancreas.
 CC -!- DOMAIN: The UEV domain is required for the interaction of the
 CC complex with ubiquitin.
 CC -!- DOMAIN: The coiled coil domain may interact with stathmin.
 CC -!- SIMILARITY: Contains 1 UEV (ubiquitin E2 variant) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U82130; AAC52083.1; -.
 DR PDB; 1KPP; NMR; A=1-145.

DR PDB; 1KPQ; NMR; A=1-145.
 DR PDB; 1M4P; NMR; A=1-145.
 DR PDB; 1M4Q; NMR; A=1-145.
 DR IntAct; Q99816; -.
 DR Genew; HGNC:15971; TSG101.
 DR H-InvDB; HIX0009490; -.
 DR MIM; 601387; -.
 DR GO; GO:0003677; F:DNA binding; TAS.
 DR GO; GO:0003714; F:transcription corepressor activity; TAS.
 DR GO; GO:0006512; P:ubiquitin cycle; TAS.
 DR InterPro; IPR010983; EF_Hand_like.
 DR InterPro; IPR008883; Tsg101.
 DR InterPro; IPR000608; UBG_conjugat.
 DR Pfam; PF05743; Tsg101; 1.
 DR PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; FALSE_NEG.
 DR PROSITE; PS50127; UBIQUITIN_CONJUGAT_2; FALSE_NEG.
 KW 3D-structure; Alternative splicing; Coiled coil; Growth regulation;
 KW Protein transport; Transport; Ubl conjugation pathway.
 FT DOMAIN 1 133 UEV.
 FT DOMAIN 235 316 Coiled coil (Potential).
 FT VARSPLIC 15 119 Missing (in isoform 2).
 FT /FTid=VSP_004440.
 FT HELIX 4 11
 FT HELIX 18 31
 FT TURN 33 34
 FT STRAND 36 44
 FT TURN 45 47
 FT STRAND 48 63
 FT TURN 64 65
 FT STRAND 66 75
 FT STRAND 86 88
 FT STRAND 103 103
 FT TURN 105 106
 FT STRAND 108 109
 FT HELIX 112 115
 FT TURN 116 116
 FT TURN 119 120
 FT HELIX 124 137
 SQ SEQUENCE 390 AA; 43944 MW; ADD6912FC22DF162 CRC64;

Query Match 100.0%; Score 2047; DB 1; Length 390;
 Best Local Similarity 100.0%; Pred. No. 2.1e-99;
 Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVSESQLKKMVS KYKYRDLTVRET VNVITLYKDLKPVLD SYVFNDGSSREL MNLTGTIP 60
 Db 1 MAVSESQLKKMVS KYKYRDLTVRET VNVITLYKDLKPVLD SYVFNDGSSREL MNLTGTIP 60
 QY 61 VPYRGNTYNIPICLWLLDTPYNPPICFVKPTSSMTIKTGKHVDANGKIYLPYLHEWKHP 120
 Db 61 VPYRGNTYNIPICLWLLDTPYNPPICFVKPTSSMTIKTGKHVDANGKIYLPYLHEWKHP 120
 QY 121 QSDLLGLIQVMIVVFGDEPPVFSRPISASYPYQATGPPNTSYMPGMPGGISPYPSGYPP 180
 Db 121 QSDLLGLIQVMIVVFGDEPPVFSRPISASYPYQATGPPNTSYMPGMPGGISPYPSGYPP 180
 QY 181 NPSGYPGCPYPPGGYPATTSSQYPSQPPVTTVGPSRDGTISED TIRASLISAVSDKLRW 240
 Db 181 NPSGYPGCPYPPGGYPATTSSQYPSQPPVTTVGPSRDGTISED TIRASLISAVSDKLRW 240
 QY 241 RMKEEMDRAQAELNALKRTEEDLKKGHQKLEEMVTRLDQEVAEVDKNIELLKKKDEELSS 300
 Db 241 RMKEEMDRAQAELNALKRTEEDLKKGHQKLEEMVTRLDQEVAEVDKNIELLKKKDEELSS 300

us-10-053-975a-1.rup

Qy	301	ALEKMQSENNDIDEVVIPTAPLYKQILNLYAEENAIEDTIFYLGEALRRGVIDLDVFL	360
Db	301	ALEKMQSENNDIDEVVIPTAPLYKQILNLYAEENAIEDTIFYLGEALRRGVIDLDVFL	360
Qy	361	KHVRLLSRKQFQLRALMQKARKTAGLSDLY	390
Db	361	KHVRLLSRKQFQLRALMQKARKTAGLSDLY	390